Comparison between J48 Decision Tree, SVM and MLP in Weather Forecasting

Suvendra Kumar Jayasingh[#1], Jibendu Kumar Mantri[*2], P. Gahan[#3]

[#1]Assistant Professor, Biju Patnaik University of Technology, India
[*2]Reader, North Orissa University, India
[#3]Professor, Sambalpur University, India

***Abstract*— *Weather forecasting is a challenging task for the Government and the general public throughout the world. Literature survey shows that the soft computing techniques play better role in predicting the weather at a particular region than the traditional mathematical or statistical methods. Now-a-days the data mining and soft computing techniques have attained the most position in research for predicting accurate weather. This paper depicts a comparison between the 3 different soft computing techniques like J48 Decision Tree, Support Vector machine and Multi Layer Perceptions (MLP) in weather forecasting. Time series data of Delhi is collected for 5 years and fed to the 3 models. After training to the 3 models, results were compared and it was concluded that the performance of J48 decision tree is consistently better.*

Keywords— *J48 Decision Tree, Support Vector Machine, Multi Layer Perceptron, Time Series Data, Weather Forecasting, WEKA (Waikato Experient and knowledge Analysis)*

I. INTRODUCTION

Weather forecasting is a major service provided by the Meteorological Department. Perfect forecasting of weather plays one vital role in day to day life of general public. Most of the agriculture in India depends on rain water and the industries also dependent on weather conditions. In case of the natural disasters and calamities, the meteorological department warns the public well before the incident which helps to be safe and take precaution for the accidents to happen. Meteorologists try for different techniques and methods for better prediction of weather. Numerical weather prediction was first proposed by Lewis Fry Rechardson in 1992. Many more soft computing techniques [10][11][12] are developed which are now used to predict the weather at a future date easily.

Our study in this paper is based upon training and testing the J48 Decion Tree, Support Vector Machine ans Multi Layer Perceptron with weather data of 5 years collected from Delhi.The objective of this study is to compare the three different models and find a conclusion to decide which model performs the best in predicting the weather at Delhi.

II. LITERATIRE SURVEY

Neeraj Kumar and Govind Kumar Jha proposed a time series ANN approach for weather forecasting in 2013[1]. Radhika, Y. and M. Shashi proposed atmospheric temperature prediction using Support Vector Machines in 2009 [2]. Paras, Sanjay Mathur have suggested a simple weather forecasting model using mathematical regression in 2012 [3]. Baboo, S. Santhosh, and I. Kadar Shereef have designed an efficient weather forecasting system using artificial neural network in 2010 [4]. Sharma, Arvind, and Manish Manoria have shown a new approach of using concept of soft computing for a weather forecasting system. Hayati, Mohsen, and Zahra Mohebi have used artificial neural network for temperature forecasting in 2007 [6]. Gill, Er Jasmeen, Er Baljeet Singh, and Er Shaminder Singh have trained back propagation neural networks with genetic algorithm for weather forecasting in 2910 [7]. Lee, Raymond, and James Liu defined a weather forecasting system using intelligent multiagent based fuzzy neuro network in 2004 [8]. Wang, Nai-Yi, and Shyi-Ming Chen have predicted temperature and TAIFEX forecasting based on automatic clustering techniques and two factors high order fuxxy time series in 2009 [9].

III. EXPERIMENTAL SETTINGS

The weather data of Delhi was collected for 5 years i.e. from January 2011 till December 2015. The weather parameters like temperature, Dew Point, humidity, Sea level Air Pressure, Wind Speed, Precipitation and events values are collected for 5 years. The data are fed to the the 3 models for training purpose then one year data was tested to check the validity of the system for use in predicting weather events. The real data may be having a lot of noise and unexpected values associated with the database. So, the data needs to be pre processed for better accuracy in prediction using the models. The WEKA (Waikato Experient and knowledge Analysis) software provides the facility to learn the 3 models by feeding the dat of 5 years and then testing it individually for validation.

IV. J48 DECISION TREE

Decision Tree is the most powerful tool in Knowledge discovery in data mining. There are many algorithms in creating the decision tree in case of data mining such as ID3, C4.5 and J48. In this paper we use J48 decision tree in building the decision tree.

V. SUPPORT VECTOR MACHINE

Support Vector Machine (SVM) is a powerful tool for data classification. It uses linear or non linear surfaces among the datasets to apply classification. The original data is mapped to feature space f with j a non linear mapping function. In this article, Support Vector Machine is used to predict the event at Delhi by analyzing the 6 other weather parameters at that particular place under discussion.

VI. MULTI LAYER PERCEPTRON

The successful application of neural network to do the data analysis is the Multi Layer Perceptron (MLP). These models are non linear neural network models which can be used for approximating a high degree of accurate prediction. It contains input layer, hidden layers and output layer.

VII. EXPERIMENTAL RESULTS

After feeding the data of 5 years to the J48 decision tree, SVM and MLP, the parameters like Correctly classified Instances, Incorrectly Classified Instances, Kappa statistic, Mean absolute error(MAE), Root mean squared error(RMSE), Relative absolute error(RAE) , Root relative squared error(RRSE) and Total Number of Instances are captured from the WEKA interface and tabulated as below for 5 distinct years.

The use of WEKA enabled us to minutely look into the different statistical parameters for 5 years of data.

The different error parameters for analyzing the prediction we have used as follows:

RSME - The root-mean-square deviation (RMSD) or root-mean-square error (RMSE) is a frequently used measure of the differences between values (sample and population values) predicted by a model or an estimator and the values actually observed.

MAE - The mean absolute error (MAE) is a quantity used to measure how close forecasts or predictions are to the eventual outcomes.

RAE - The relative error is the absolute error divided by the magnitude of the exact value. The percent error is the relative error expressed in terms of per 100.

RRSE - The Root relative squared error is calculated as the Mean absolute error divided by the error of the ZeroR classifier (a classifier, that ignores all predictors and simply selects the most frequent value).

Table I. Parameters i.e Correctly classified Instances, Incorrectly Classified Instances, Kappa statistic, Mean absolute error(MAE), Root mean squared error(RMSE), Relative absolute error(RAE) , Root relative squared error(RRSE) and Total Number of Instances are captured from the WEKA for 5 different years (2011 to 2015)

2011				
Sl. No.	Parameters	J48 Decision Tree	SVM	MLP
1	Correctly classified Instances	289(79.1781%)	238(65.2055%)	276(75.6164%)
2	Incorrectly Classified Instances	76(20.8219%)	127(34.7945%)	89(24.3836%)
3	Kappa statistic	0.6624	0.4141	0.6109
4	Mean absolute error(MAE)	0.0969	0.212	0.0965
5	Root mean squared error(RMSE)	0.2201	0.314	0.2233
6	Relative absolute error(RAE)	49.83%	109.06%	49.65%
7	Root relative squared error(RRSE)	70.76%	1009.97%	71.79%
8	Total Number of Instances	365	365	365

2012				
Sl. No.	Parameters	J48 Decision Tree	SVM	MLP
1	Correctly classified Instances	289(78.9617%)	252(68.8525%)	280(76.5027%)
2	Incorrectly Classified Instances	77(21.0383%)	114(31.1475%)	86(23.4973%)
3	Kappa statistic	0.6173	0.3977	0.5755
4	Mean absolute error(MAE)	0.0793	0.1919	0.0843
5	Root mean squared error(RMSE)	0.1991	0.2988	0.2119
6	Relative absolute error(RAE)	51.64%	125.06%	54.95%
7	Root relative squared error(RRSE)	72.19%	108.34%	76.83%
8	Total Number of Instances	366	366	366

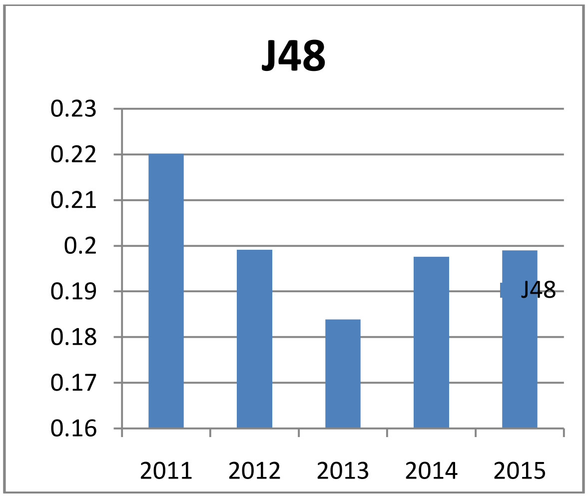
2013				
Sl. No.	Parameters	J48 Decision Tree	SVM	MLP
1	Correctly classified Instances	300(82.1918%)	244(66.8493%)	282(77.2603%)
2	Incorrectly Classified Instances	65(17.8082%)	121(33.1507%)	83(22.7397%)
3	Kappa statistic	0.7061	0.4104	0.6185
4	Mean absolute error(MAE)	0.0677	0.1932	0.077
5	Root mean squared error(RMSE)	0.1839	0.3008	0.2016
6	Relative absolute error(RAE)	41.97%	119.88%	47.74%
7	Root relative squared error(RRSE)	65.04%	106.35%	71.30%
8	Total Number of Instances	365	365	365

2014				
Sl. No.	Parameters	J48 Decision Tree	SVM	MLP
1	Correctly classified Instances	289(79.1781%)	252(69.0411%)	274(75.0685%)
2	Incorrectly Classified Instances	76(20.8219%)	113(30.9589%)	91(24.9315%)
3	Kappa statistic	0.649	0.4247	0.5596
4	Mean absolute error(MAE)	0.0781	0.1921	0.0861
5	Root mean squared error(RMSE)	0.1976	0.2991	0.2115
6	Relative absolute error(RAE)	49.43%	121.56%	54.50%
7	Root relative squared error(RRSE)	70.60%	106.85%	75.56%
8	Total Number of Instances	365	365	365

2015				
Sl. No.	Parameters	J48 Decision Tree	SVM	MLP
1	Correctly classified Instances	303(83.0137%)	270(73.9726%)	294(80.5479%)
2	Incorrectly Classified Instances	62(16.9863%)	95(26.0274%)	71(19.4521%)
3	Kappa statistic	0.7164	0.5351	0.6712
4	Mean absolute error(MAE)	0.0792	0.2105	0.0783
5	Root mean squared error(RMSE)	0.199	0.3116	0.2073
6	Relative absolute error(RAE)	43.52%	115.73%	43.04%
7	Root relative squared error(RRSE)	66.19%	103.64%	68.95%
8	Total Number of Instances	365	365	365

VIII. COMPARISON OF PARAMETERS

Now let's compare the Root mean squared error(RMSE), Mean absolute error(MAE), Relative absolute error(RAE), Root relative squared error(RRSE) and Root relative squared error(RRSE) for data of 5 years between J48 Decision tree, SVM and MLP.



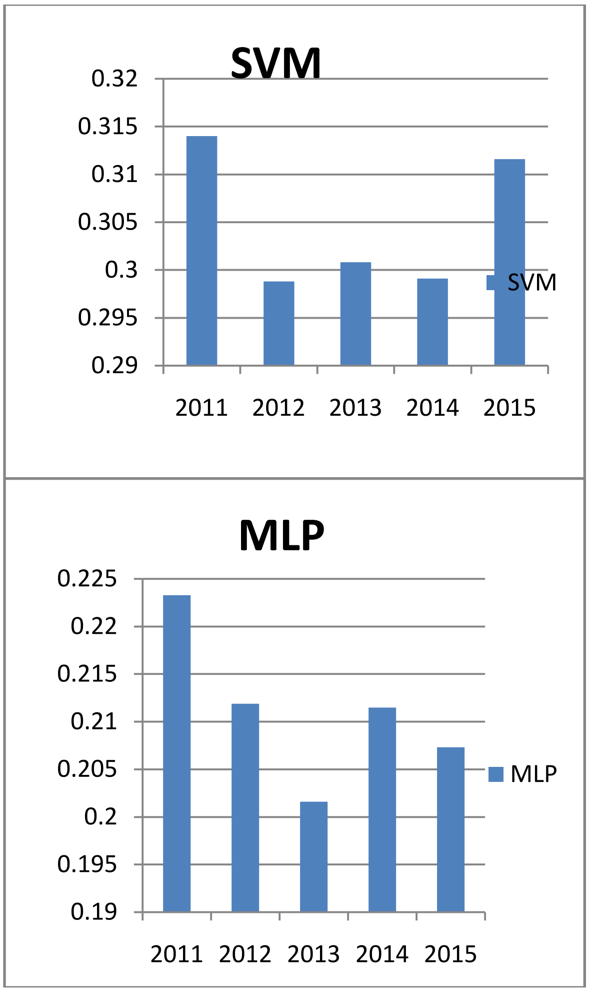


FIGURE 1. The Comparison of Root mean squared error (RMSE) between J48, SVM and MLP

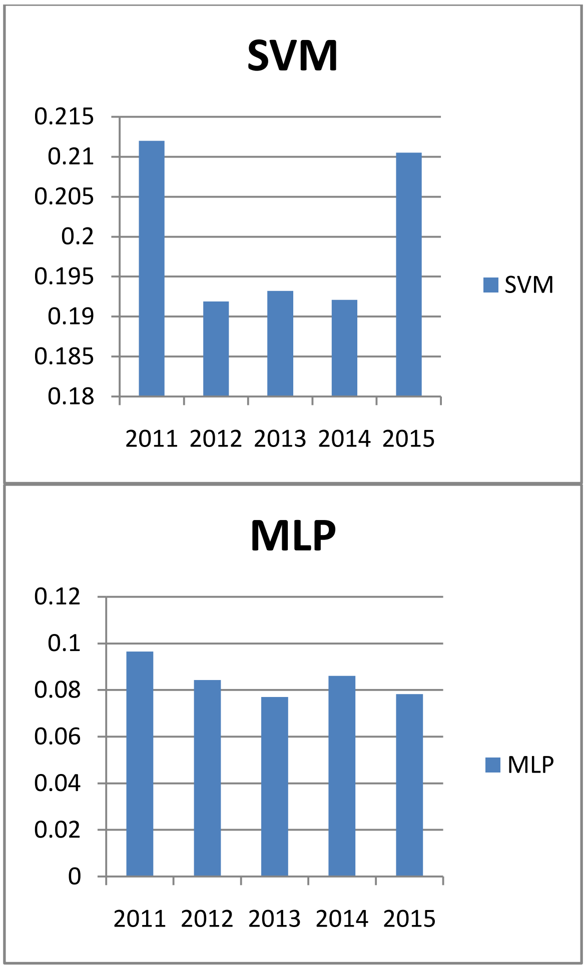
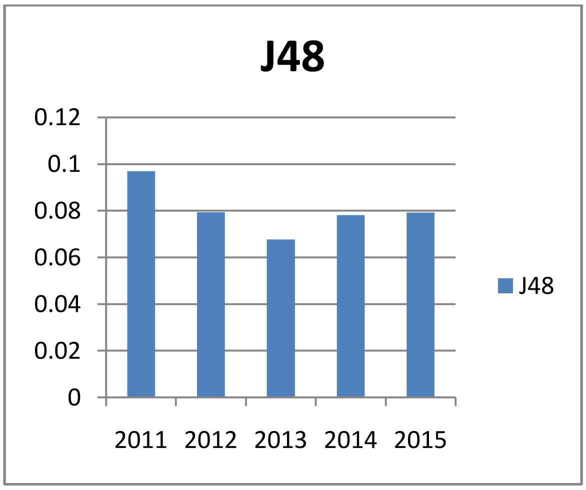
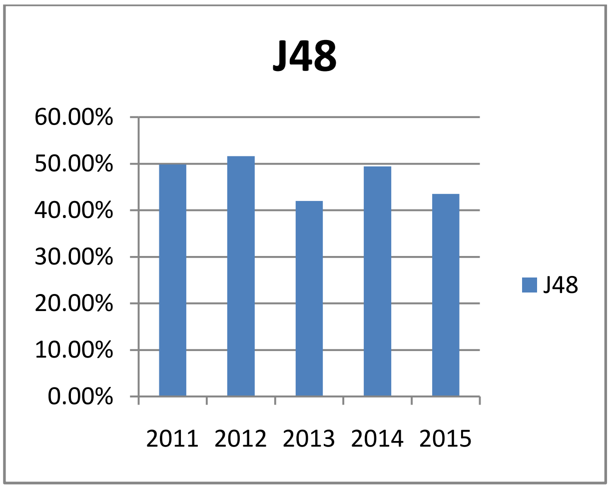


FIGURE 2. The Comparison of Mean Absolute Error(MAE) between J48, SVM and MLP





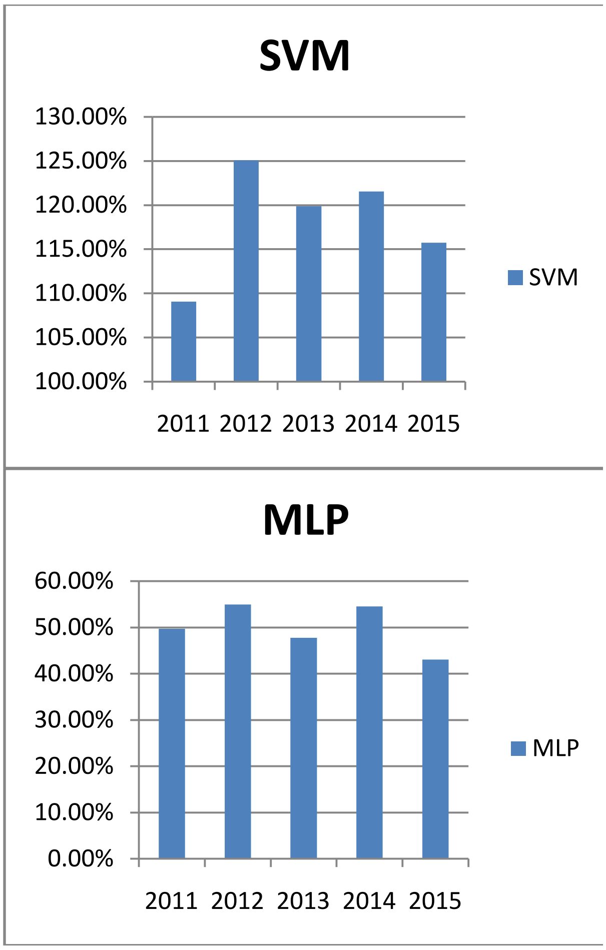
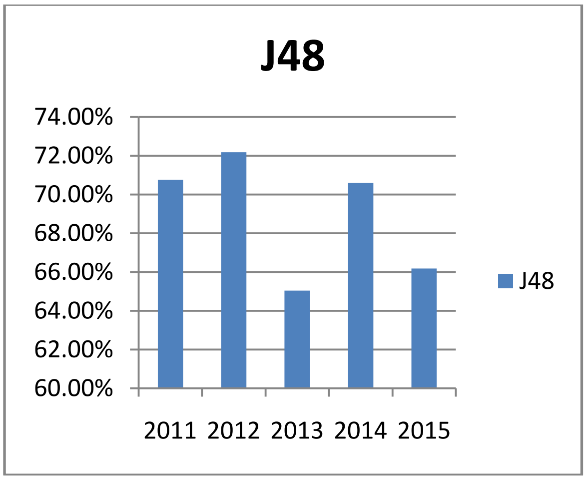


FIGURE 3. The Comparison of Relative Absolute Error(RAE) between J48, SVM and MLP



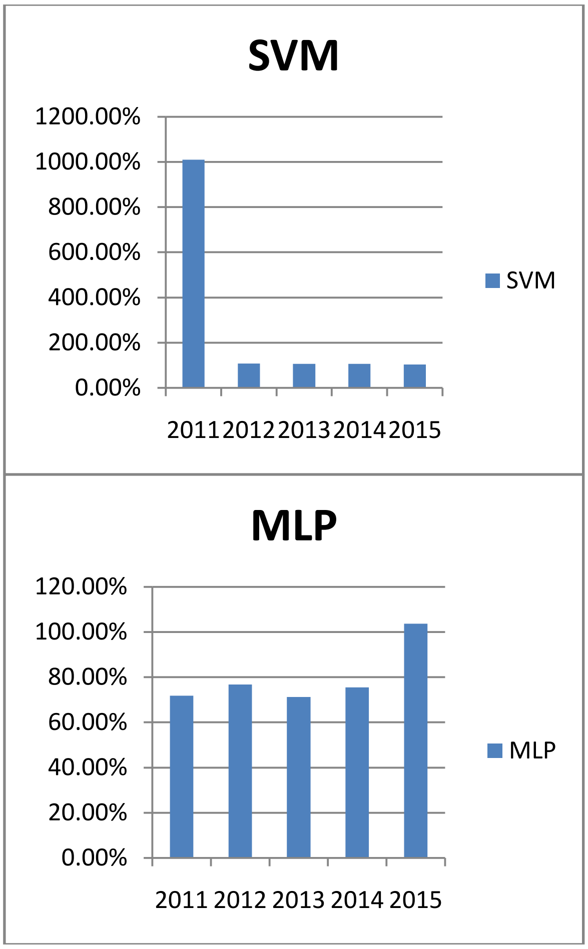


FIGURE 4. The Comparison of Root Relative Squared Error(RRSE) between J48, SVM and MLP

IX. CONCLUSION AND FUTURE WORK

From the analysis of 4(Four) different error parameters among the 3(Three) different models for the weather data of 5(Five) years at Delhi, it is concluded that J48 decision tree performs consistently better than the other 2(two) for accurate and better weather forecasting. So, in future the model of J48 decision tree may be enhanced to better forecast the rain fall, cyclone, storm and other natural disaster and calamities in near future thereby saving the life of lakhs of human beings and domestic animals. Better prediction of natural disasters and calamities will make the people aware to safeguard themselves. The forecasting of weather with better accuracy and less prediction error is really useful in the field of agriculture, mountaineering, fishing in sea and many more day to day activities of human being. The better the prediction, the safer will be the people and the properties. The J48 decision tree may be used for many more such forecasting problems in future.

`

X. REFERNCES

[1] Neeraj Kumar , Govind Kumar Jha "A Time Series ANN Approach for Weather Forecasting" International Journal of Control Theory and Computer Modeling (IJCTCM) January 2013

[2] Radhika, Y., and M. Shashi. "Atmospheric temperature prediction using support vector machines." *International Journal of Computer Theory and Engineering* 1.1 (2009): 55.

[3] Paras, Sanjay Mathur. "A Simple Weather Forecasting Model Using Mathematical Regression." *Indian Research Journal of Extension Education*12.2 (2016): 161-168.

[4] Baboo, S. Santhosh, and I. Kadar Shereef. "An efficient weather forecasting system using artificial neural network." *International journal of environmental science and development* 1.4 (2010): 321 - 326.

[5] Sharma, Arvind, and Manish Manoria. "A Weather Forecasting System using concept of Soft Computing: A new approach." *2006 International Conference on Advanced Computing and Communications*. IEEE, 2006.

[6] Hayati, Mohsen, and Zahra Mohebi. "Application of artificial neural networks for temperature forecasting." *World Academy of Science, Engineering and Technology* 28.2 (2007): 275-279.

[7] Gill, Er Jasmeen, Er Baljeet Singh, and Er Shaminder Singh. "Training back propagation neural networks with genetic algorithm for weather forecasting."*IEEE 8th International Symposium on Intelligent Systems and Informatics*. IEEE, 2010.

[8] Lee, Raymond, and James Liu. "iJADE WeatherMAN: a weather forecasting system using intelligent multiagent-based fuzzy neuro network." *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* 34.3 (2004): 369-377.

[9] Wang, Nai-Yi, and Shyi-Ming Chen. "Temperature prediction and TAIFEX forecasting based on automatic clustering techniques and two-factors high-order fuzzy time series." *Expert Systems with Applications* 36.2 (2009): 2143-2154.

[10] Sharma M., Mathew L. And Chatterji S., “Weather Forecasting using Soft Computing and Statistical Techniques”. *International Journal of Advanced Research in Electrical, Electronics and Instrumental Engineering*, ISSN ONLINE (2278-8875) PRINT(2320-3765), 2016.

[11] Rahul G.K., Khurana M. And Sinha A.K., “Weather Forecasting Using Soft Computing: Minimum Temperature, Maximum Temperature & Pressure”, *International Journal of Engineering Research and Developmemt*. Volumeb7, Issue 11(July 2013), pp. 10-17.

[12] Bhatkande S. S. And Hubballi R.G., “Weather Predictionbased on Decision Tree Algorithm Using Data Mining Techniques”, *International Journal of Advanced Research in Computer and Communication Engineering*, Vol. 5, Issue 5, May 2016.